Improving Estimation of Wolf Recruitment and Abundance, and Development of an Adaptive Harvest Management Program for Wolves in Montana

Federal Aid in Wildlife Restoration Grant W-161-R-1
Annual interim report, March 2018

Sarah Sells
PhD Candidate
Montana Cooperative Wildlife Research Unit
205 Natural Sciences, Missoula, MT 59812

Justin Gude
Res & Tech Services Bureau Chief
Montana Fish, Wildlife and Parks
1420 E. 6th St., Helena, MT 59620

Allison Keever
PhD Candidate
Montana Cooperative Wildlife Research Unit
205 Natural Sciences, Missoula, MT 59812

Kevin Podruzny
Biometrician
Montana Fish, Wildlife and Parks
1420 E. 6th St., Helena, MT 59620

Mike Mitchell
Unit Leader
Montana Cooperative Wildlife Research Unit
205 Natural Sciences, Missoula, MT 59812

State: Montana
Agency: Fish, Wildlife & Parks
Grant: Montana wolf monitoring study
Grant number: W-161-R-1
Time period: January 1, 2017–December 31, 2017
INTRODUCTION

Wolves (*Canis lupus*) were reintroduced in the northern Rocky Mountains (NRM) in 1995, and after rapid population growth were delisted from the endangered species list in 2011. Since that time, states in the NRM have agreed to maintain populations and breeding pairs (a male and female wolf with 2 surviving pups by December 31; USFWS 1994) above established minimums (≥150 wolves and ≥15 breeding pairs within each state). Montana estimates population size every year using patch occupancy models (POM; MacKenzie et al. 2002, Rich et al. 2013, Miller et al. 2013, Bradley et al. 2015), however, these estimates are sensitive to pack size and territory size, and were developed pre-harvest. Reliability of future estimates based on POM will be contingent on accurate information on territory size, overlap, and pack size, which are expected to be strongly affected by harvest. Additionally, breeding pairs, which has proven to be an ineffective measure of recruitment, are determined via direct counts. Federal funding for wolf monitoring has ended in states where wolves are delisted, and future monitoring will not be able to rely on intensive counts of the wolf population. Furthermore, intensive, field-based monitoring has become cumbersome and less effective since the population has grown. With the implementation of harvest, it is pertinent to predict the effects of harvest on the wolf population and continue to monitor to determine effectiveness of management actions to make informed decisions regarding hunting and trapping seasons.

STUDY OBJECTIVES

Our 4 study objectives are to:

1. Improve estimation of recruitment.

2. Improve and maintain calibration of wolf abundance estimates generated through POM.

3. Develop a framework for dynamic, adaptive harvest management based on achievement of objectives 1 & 2.

4. Design a targeted monitoring program to provide information needed for robust estimates and reduce uncertainty in the AHM paradigm over time.

Two PhD students are addressing the 4 study objectives as part of Project 1 (Sarah Sells) and Project 2 (Allison Keever; Fig. 1).

DELIVERABLES

1. A method to estimate recruitment for Montana’s wolf population that is more cost effective and biologically sound than the breeding pair metric (Project 2, A. Keever).
2. Models to estimate territory size and pack size that can keep POM estimates calibrated to changing environmental and management conditions for wolves in Montana (Project 1, S. Sells).

3. An adaptive harvest management model that allows the formal assessment of various harvest regimes and reduces uncertainty over time to facilitate adaptive management of wolves (Project 2, A. Keever).

4. A recommended monitoring program for wolves to maintain calibration of POM estimates, determine effectiveness of management actions, and facilitate learning in an adaptive framework (Projects 1 & 2).

LOCATION

This study encompasses wolf distribution in Montana and Idaho (Fig. 2). Additional data will come from Yellowstone National Park for the territory models developed under objective 2.
GENERAL PROGRESS

Projects 1 & 2, Year 1:
We (S. Sells & A. Keever) started our PhD programs in January 2015 (Fig. 3).

Much of year 1 was devoted to literature reviews on animal behavior, carnivores, modeling, optimal foraging, etc. and determining approaches for the dissertations. We also formed and held multiple meetings with our committees, worked on completing coursework requirements, and finalized research statements. Additional efforts focused on communicating with wolf specialists, identifying target packs for collaring, managing collar orders and data, and helping coordinate contracts and capture plans for winter aerial captures for January and February 2016. We also met with wolf specialists in the field to learn more about the wolves in each region, and coordinated and held meetings with the specialists to plan future project efforts.

Project 1 (S. Sells): In year 2, I continued most activities from year 1, including conducting literature searches, holding committee meetings, communicating with wolf specialists, managing collar orders, managing data, etc. I also began working on the theoretical territory models. My primary focus was meeting project and university requirements and deadlines, including defending my proposal and passing my comprehensive exams. I also joined the wolf specialists to assist with a month of trapping.

Year 3 was primarily devoted to preparing the theoretical territory models. I presented draft results at 5 conferences. In addition to completing more coursework, I continued working with MFWP and collar manufacturers as the point person on ordering collars, troubleshooting a growing set of issues with the collars, and managing collar records. I continued coordinating data management and collection from deployed collars and communicating with wolf specialists on all trapping and collar-related topics. I also spent 2 weeks assisting wolf specialists with trapping.

Project 2 (A. Keever): In year 2 I continued literature reviews, completed coursework, and meeting university requirements. I defended my proposal and was studying for my comprehensive exams. Another focus was on the empirical recruitment model. I began developing the model that I had outlined in my proposal. I also spent 1 month assisting wolf specialists with trapping.
Year 3 I completed the empirical recruitment model code and tested the model with simulated data. Much of my time was spent compiling and formatting the data needed to estimate recruitment. I presented preliminary results at 2 conferences. I also passed my comprehensive exams and spent 2 weeks assisting wolf specialists with trapping.

**Deliverables and updates:** Project deliverables will include an empirical recruitment model; theoretical territory, group size, and recruitment models; draft and final AHM models; and final territory and pack size models. We have been working on deliverables of the empirical recruitment model (A. Keever) and the theoretical territory models (S. Sells) towards meeting objectives 1 and 2. We each describe our progress towards these deliverables in this report. (Additional details on objectives 3 and 4 are available in the 2016 report.)

**DATA COLLECTION SUMMARY**

Trapping efforts by Montana Fish, Wildlife and Parks have continued since 2014:
- There have been 66 successful captures directly related to this project through 2017.
- Collars were deployed in approximately 46 packs (this number is fluid as wolves disperse).
- Using ground and aerial captures:
  - 10 collars were deployed in 2014.
  - 14 collars were deployed in 2015.
  - 27 collars were deployed in 2016.
  - 16 collars were deployed in 2017.
- These collars have yielded >26,000 locations of wolves (Fig. 4).
- After collar removals, harvests, other mortalities, and some collar losses (e.g., through dropped collars), 28 collars remained deployed at the end of 2017.
- Many of the collars began experiencing major performance issues in 2017, however. Of the 28 deployed collars, only 9 were functional as of December 2017 (see below). Collaring efforts will continue via ground and aerial captures through 2018.

The project began experiencing a growing set of technical issues with collars in 2017. Many collars began failing to send reliable transmissions to the satellite service, and eventually many stopped transmitting altogether. After 3 months without a fix, a collar is considered to have malfunctioned and is deactivated. In summary:
- 20 collars have failed while deployed, 16 of which are still deployed.
  - 15 collars worked for 1 – 2.25 years before failing.
    - 1 collar was recovered and had VHF failure.
    - 2 collars were recovered and had battery failure.
    - 12 collars are still deployed.
  - 4 collars worked for <1 year (4 – 11 months) before failing and are still deployed.
  - 1 collar never worked after deployment and was recently recovered.
- 3 more collars are approaching the 3 month deadline without a transmission and will be deactivated soon.
- 9 collars are functional or mostly functional as of December 2017.

We are working with Lotek to return all collars that have not yet been deployed. These will be replaced with collars that are expected to provide better performance based on what Lotek has learned from these recent failures. We consider these collar failures and challenges all the more impetus to reduce needs for future collaring efforts; our work will help achieve this goal.

Fig. 4. Locations of wolves collared for this project, 2014−2016. Colors represent different wolves. Note that some polygons include dispersal from original pack’s territory.
PROGRESS ON OBJECTIVES

Objective 1: Improve estimation of recruitment—Allison Keever, Project 2

1.1 Background

Estimating recruitment (i.e., number of young produced that survive to an age at which they contribute to the population) of wolves can be difficult due to their complex social structure. Wolves are cooperative breeders, and pack dynamics (e.g., pack tenure, breeder turnover, and number of non-breeding helpers) can affect recruitment and pup survival (e.g., Ausband et al. 2015). Cooperative breeding often relies on the presence of non-breeding individuals that help raise offspring (Solomon and French 1997), and reduction in group size can lead to decreased recruitment in cooperative breeders (Sparkman et al. 2011, Stahler et al. 2013). Human-caused mortality through both direct and indirect means (Ausband et al. 2015) and prey biomass per wolf (Boertje and Stephenson 1992) have been shown to affect recruitment. As a result, it will be important to consider the effects of harvest, pack dynamics, wolf density, and prey availability on recruitment.

Further challenges of estimating recruitment include the size of the wolf population and limited time and funding for monitoring. Currently, MFWP documents recruitment through visual counts of breeding pairs (a male and female wolf with 2 surviving pups by December 31; U.S. Fish and Wildlife Service 1994). These counts, however, are likely incomplete due to the large number of wolves in the population. Federal funding for wolf monitoring in Montana and Idaho is no longer available. States therefore fund their own monitoring programs, and future monitoring will not be able to rely on intensive counts. A breeding pair estimator (Mitchell et al. 2008) could be used to estimate breeding pairs, but this requires knowing pack size; such data are hard to collect given the size of the wolf population. Additionally, the breeding pair metric is an ineffective measure of recruitment because it provides little insight into population growth rate or the level of harvest that could be sustained. Recruitment could be estimated by comparing visual counts at the den site to winter counts via aerial telemetry (Mech et al. 1998) or by marking pups at den sites (Mills et al. 2008). An alternative method could include non-invasive genetic sampling (Ausband et al. 2015) at predicted rendezvous sites (Ausband et al. 2010). These methods, however, may not be feasible on large scales due to budget and staff constraints. Existing monitoring efforts yield insufficient data to estimate recruitment using traditional methods; therefore a new approach is needed that does not rely on extensive data.

1.2 Goals and General Approach

Our objective is to develop an approach to estimate recruitment that is more tractable, cost effective, and biologically credible than the breeding pair metric. Collar and count data are currently collected for on-going monitoring, however these data may not be available or at least not as many data available moving forward. Therefore, our goal is to create a model that can be
flexible in the amount of data required to estimate recruitment and also evaluate the accuracy and precision of estimates with varying amounts of data. Integrated population models can be a useful tool for demographic analyses from limited data sets, and can increase precision in estimates (Besbeas et al. 2002). We will develop a per capita integrated population model (hereafter IPM) to estimate recruitment and evaluate the relationship between recruitment and factors that may cause spatial and temporal variation in wolf recruitment. We will use collar, count and hunter survey data from 2007–2016 in Montana to estimate recruitment. We will also use a simulation study to evaluate how many data are needed to get reliable estimates using this method to see if it will be cost effective to implement.

The resulting statistical model will relate covariates and recruitment. It will not, however, improve understanding of the mechanisms that cause recruitment to change. Recruitment depends on a pack’s success in breeding and giving birth, as well as litter size and pup survival. Whether a pack successfully breeds and gives birth or not is primarily determined by the survival of the breeding pair in the pack. Conversely, pup survival may be affected by helper presence, prey availability, disease outbreaks, and human-caused mortality (Goyal et al. 1986, Boertje and Stephenson 1992, Johnson et al. 1994, Mech and Goyal 1995, Fuller et al. 2003, Ausband et al. 2015). Unfortunately, there are few data to estimate the contribution of those factors to overall pup recruitment, so we will also develop a mechanistic model of recruitment to theoretically explore the effects of human-caused mortality, prey availability, multiple litters per pack, disease outbreaks, and group size on the different components of recruitment. The probability a pack successfully breeds and reproduces, litter size per pack, and pup survival all determine pup recruitment. Hypotheses about how factors such as disease, harvest, or prey availability affect these parameters can be explored using liner or non-linear models and then multiplied together. Different models can be developed that represent different hypotheses. Those different hypotheses will result in different predictions of recruitment if those hypotheses were correct. The model predictions can be compared to estimated recruitment from the IPM to determine which hypotheses have most support.

1.3 Methods

We are currently developing the IPM model to estimate recruitment in program R (R Core Team 2014) in a Bayesian framework using package R2jags (Su and Yajima 2015) to communicate with JAGS (Plummer 2003). The IPM model will allow us to evaluate the factors that cause spatial and temporal variation in recruitment and, through use of a simulation study, determine data requirements for estimating recruitment. Recruitment data are not available across Montana, so we will use the hunter survey, group count, and GPS and VHF collar data that are currently available from ongoing monitoring. The IPM will have a 1) POM model to estimate abundance, 2) survival model, 3) recruitment model, 4) a population-level model to relate changes in abundance over time with survival and recruitment, and 5) a group-level model to relate changes in group size over time with survival and recruitment (Fig. 1.1). This IPM framework is unique
in that it adds a group-level model to account for the social structure of wolves and its influence on recruitment. We are evaluating the efficacy of the IPM model by simulating data to test how many data are required for accurate estimates of recruitment. Then, we will use hunter survey, group count, and collar data to estimate recruitment across the state of Montana.

**1.3.1 POM model**

We will use the same occupancy modeling framework that MFWP currently applies across the state using hunter survey data to estimate abundance of the wolf population. We will use a dynamic false-positive occupancy model (MacKenzie et al. 2002; Miller et al. 2013, Rich et al. 2013; Bradley et al. 2015) to estimate the area occupied by wolves. We will then use GPS collar data from 2008-2009 (Rich et al. 2012) to estimate mean territory size. The number of packs is then the area occupied by wolves divided by the mean territory size. To estimate abundance we will take the number of packs and multiply by the average group size of a wolf pack. Group size will be estimated from the group count data while accounting for observation error for each year. We will also account for territory overlap like MFWP does for their abundance estimates. Eventually, work from current research (Objective 2) on territory size and group size will be used in place of average territory and group size to improve estimates of abundance in the IPM model.

**1.3.2 Survival model**

We will estimate survival using a discrete-time proportional hazards model, or a complementary log-log (cloglog) model. We will use biologically relevant discrete periods for analyses such as
the denning period (April-May), rendezvous period (June-August), and the hunting/trapping season (September-March). GPS and VHF collared wolves from 2007-2016 will provide the known-fate information needed to estimate survival. These data, however, may have inherent sampling bias. Most collared wolves from this time period are targeted because they are livestock conflict packs. These data would bias survival low. To account for this we could use an informative prior on survival and weight the collars so that research collars have more influence on the posterior estimate of survival than collars from livestock conflict packs. Or, we could also only use collars deployed for research purposes to account for this bias in survival.

1.3.3 Recruitment model

We will evaluate factors that explain the spatial and temporal variation in recruitment using generalized linear models with a log link function. We will develop *a priori* hypotheses regarding how factors such as human-caused mortality rates, landowner-type (e.g., public vs. private), road density, land cover type, elevation, and group size affect recruitment of wolves. We will test these hypotheses using the IPM in Montana.

1.3.4 Population level

Changes in abundance over time are a function of births, deaths, immigration, and emigration. We have information about abundance and survival for wolves in Montana, therefore we can essentially solve for recruitment. Because the pack is the reproductive unit, at the population level we will account for immigration and emigration by including colonization and extinction of packs which will be informed by the occupancy model. Lone wolves that immigrate into the population can be ignored. Wolves joining or dispersing from a pack will be accounted for at the group level.

1.3.5 Group-level model

A typical IPM framework does not account for animals with social structure and cooperative breeding. Therefore, we will add a group level model that explicitly accounts for the social structure of wolves. This framework allows us to estimate recruitment at the pack level as well as the population level which improves estimation of recruitment. We will also include dispersal from the pack modeled using recent literature on dispersal rates of wolves in the U.S. northern Rocky Mountains (Jimenez et al. 2017). Changes in group size will be a result of recruitment, survival, and dispersal.

The main objective of this work is to provide a method to estimate recruitment that is more cost effective, which means it cannot require a lot of data. This framework requires group count data to estimate recruitment. These data, however, may be too costly to collect in the future. The IPM is flexible and could still estimate recruitment with only the population-level. Therefore, we will test the IPM without the group-level as well which would eliminate the need for group count data.
1.3.6 Data simulation

Our goal is to provide a model to estimate recruitment that is more cost effective. For a method to be cost effective, and therefore useful for monitoring, it cannot rely on a lot of data that are expensive to collect. To determine whether the IPM model would be useful in the future we evaluated the amount of data that would be needed to get reliable estimates of recruitment using a simulation study. We simulated a wolf population for 10 years and then sampled from the population. To do this we first generated 100 wolf packs using a Poisson distribution with an average pack size of 4 wolves. We then randomly generated survival, recruitment, and dispersal rates using a uniform distribution with a range of biologically realistic rates. This allowed for yearly variation in the demographic rates, which we could then record as our truth. The simulated wolves then survived and reproduced based on the demographic rates we generated, with stochasticity using Poisson and binomial distributions for reproduction and survival/dispersal, respectively. We then added up the number of wolves within packs to get truth for total abundance.

After simulating the wolf population, we then randomly sampled 50, 25, and 12 packs of the 100 for group count data. We also added observation error, so our sample of packs is also a sample of wolves within the pack. For survival data we used our truth survival for each year and generated 50 known-fate observations of wolves incorporating stochasticity using the binomial distribution. We then sampled 20, 10, and 5 of those observations which represent our collar data. We used these data in the IPM model to estimate recruitment and determine how well it matched our truth we used to simulate the data.

1.4 Preliminary Results

With simulated data we know “truth,” and can compare our estimates to truth. We ran the IPM model with occupancy fixed to evaluate the amount collar and group count data needed for accurate estimates of recruitment. We also compared our estimates of survival, group size, and abundance to truth. We found that datasets with at least 10 collars and 25 group counts were precise for estimating recruitment (Fig. 1.2). Generally, all datasets except the dataset with 5 collars and 12 group counts provided accurate estimates of recruitment with a % error of < 20% (Table 1.1). All datasets provided approximately the same accuracy of abundance estimates (Fig. 1.3), and only the dataset with 5 collars resulted in inaccurate estimates of survival (Fig. 1.4).
Fig. 1.2: Estimates of recruitment (pups per pack) generated from varying amounts of data compared to truth: full dataset (50 group counts; 20 collars), A dataset (25 group counts; 20 collars), B dataset (50 group counts; 10 collars), C dataset (25 group counts; 10 collars), and D dataset (12 group counts; 5 collars).

Fig. 1.3: Estimates of abundance generated from varying amounts of data compared to truth: full dataset (50 group counts; 20 collars), A dataset (25 group counts; 20 collars), B dataset (50 group counts; 10 collars), C dataset (25 group counts; 10 collars), and D dataset (12 group counts; 5 collars).

Table 1.1: % error of recruitment estimates from truth from varying amounts of data: full dataset (50 group counts; 20 collars), A dataset (25 group counts; 20 collars), B dataset (50 group counts; 10 collars), C dataset (25 group counts; 10 collars), and D dataset (12 group counts; 5 collars).

<table>
<thead>
<tr>
<th></th>
<th>Full dataset</th>
<th>A dataset</th>
<th>B dataset</th>
<th>C dataset</th>
<th>D dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error (%)</td>
<td>13.8%</td>
<td>18.0%</td>
<td>14.2%</td>
<td>19.4%</td>
<td>38.5%</td>
</tr>
</tbody>
</table>
1.5 Summary and next steps

The objective of this work is to provide a method to estimate recruitment that is both biologically credible and cost effective. The main determinant of whether this method will be cost effective is the amount of data required to estimate recruitment. The IPM can be a viable method to estimate recruitment because reliable estimates are generated using only 12-25 group counts and 10 collars. Further, if group count data are too costly to collect, the model can be adjusted to eliminate the need for group count data by removing the group-level model. Our next step will be to test the model without the group-level and evaluate the accuracy of recruitment estimates. The tradeoff between resources spent collecting data and accuracy of estimates generated from those data can then be assessed.

The other objective of this work was to provide a method that is more biologically credible than the breeding pair metric. The breeding pair metric estimates the probability a pack contains a breeding pair and does not provide detailed information on recruitment. The IPM model, which has been developed to account for wolves’ social structure, is a method that provides accurate estimates of recruitment that we can use to answer biological questions about spatial and temporal variation in recruitment. This information can then be used to help inform harvest decisions.

We have recently completed data formatting and will begin running models to estimate recruitment of wolves across Montana. We will test a priori hypotheses about the factors that cause spatial and temporal variation in recruitment and use model selection to determine which hypotheses have most support. Then, we will apply the model in Idaho using the same types of data sources and compare model estimates of recruitment with field-based recruitment data as an external test of the model.
Objective 2: Improve and maintain calibration of wolf abundance estimates generated through POM—Sarah Sells, Project 1

2.1 Introduction

Monitoring is a critical, yet challenging, management tool for gray wolves. Since delisting of wolves in 2011, monitoring results help MFWP set management objectives and communicate with stakeholders and the public. Monitoring any large carnivore is challenging, however, due to their elusive nature and naturally low densities (Boitani et al. 2012). This is particularly true for wolves due to increasing populations, decreasing funding for monitoring, and changing behavioral dynamics with harvest.

Abundance estimates are a key component of monitoring (Bradley et al. 2015). Abundance is currently estimated in Montana with 3 parameters: area occupied, average territory size, and annual average pack size (Fig. 2.1, Bradley et al. 2015). Area occupied is estimated with a Patch Occupancy Model (POM) based on hunter observations and field surveys (Miller et al. 2013, Bradley et al. 2015). Average territory size is assumed to be 600 km² with minimal overlap, based on past work (Rich et al. 2012). Annual average pack size is estimated from monitoring results. Total abundance (N) is then calculated as: \( N = \text{area occupied} / \bar{x} \text{ territory size} \times \bar{x} \text{ pack}. \)

Whereas estimates of area occupied from POM are expected to be reliable (Miller et al. 2013, Bradley et al. 2015), reliability of abundance estimates hinge on key assumptions about territory size, territory overlap, and pack size (Bradley et al. 2015). Assumptions of fixed territory size and minimal overlap are simplistic; in reality, territories vary spatiotemporally (Uboni et al. 2015). This variability is likely even greater under harvest (Brainerd et al. 2008). Meanwhile, pack size estimates assume all packs are located and accurately counted each year, which is no longer possible due to the number of packs and declining funding for monitoring (Bradley et al. 2015). Since implementation of harvest in 2009, several factors have further compounded these challenges and decreased accuracy of pack size estimates. First, whereas larger packs are generally easier to find and monitor, average pack size has decreased since harvest began (Bradley et al. 2015). Difficult-to-detect smaller packs may be more likely to be missed altogether,
biasing estimates of average pack size high. Conversely, incomplete pack counts, especially for larger packs, could bias estimates of average pack size low. Harvest and depredation removals also affect social and dispersal behavior (Adams et al. 2008, Brainerd et al. 2008, Ausband 2015). Additionally, pack turnover is now greater than in populations with less human-caused mortality.

Development of reliable methods to estimate territory size, territory overlap, and pack size is critical for accurate estimates of abundance. One means for developing models to estimate territories and pack sizes is an empirical modeling approach. This approach generally involves measuring and attempting to discern patterns in territory and pack size dynamics (e.g., Rich et al. 2012). Empirical models do not, however, provide an understanding of causal mechanisms, i.e., the underlying processes that shape the system and patterns we observe, such as processes driving decisions carnivores make about where to settle and whether to stay in or leave a social group. Ignoring causal mechanisms may yield models that do not suitably predict conditions beyond the spatiotemporal scale for which they were developed (Mitchell and Powell 2002). Empirical models may also require extensive continued monitoring and data collection to provide sufficient data for predictions.

An alternative method to empirical modeling is a mechanistic modeling approach. Such an approach involves developing theoretical models that capture the hypothesized causal mechanisms structuring the system (Mitchell & Powell 2004, 2012). Mechanistic models may take the form of individual-based models (IBMs, also known as agent-based models). Although often challenging to develop, IBMs provide an ideal means for understanding the mechanisms driving territorial behavior. Consistent with the role of individuals in natural selection (Darwin 1859), IBMs are bottom-up whereby population-level behaviors and patterns emerge from the interaction of individuals with one-another and their environment (Grimm and Railsback 2005, Grimm et al. 2005, DeAngelis and Grimm 2014). IBMs therefore differ strongly from traditional population models that rely on differential equations and impose top-down population parameters (e.g., birth rate; DeAngelis and Grimm 2014). As a result, IBMs are less abstract and easier to conceptualize. Once designed, IBMs offer “virtual laboratories” for investigating how bottom-up influences of individuals give rise to complex organization of the larger system (Grimm et al. 2005). Predictions from these models can be compared to actual behaviors of animals to identify the model(s) with most support (Mitchell & Powell 2002, 2004, 2007, 2012). Resulting models are based on the likely causal mechanisms that shape the system, and thus yield reliable scientific inference and are predictive at any spatiotemporal scale. Importantly, abundant data are not required for predictions once models are developed.

2.2 Goals and General Approach

Our goal is to develop tools to estimate territory and group size of wolves to calibrate estimates of abundance of wolves from POM in Montana and Idaho. To achieve this goal, our steps will be to:
1. **Develop a suite of mechanistic territory models.** These models will capture the potential causal mechanisms we hypothesize structure territories of wolves. We will run simulations to provide general predictions of territorial behavior under each model.

2. **Identify the most predictive territory model for wolves in Montana and Idaho.** We will summarize general patterns of territories (e.g., their size and overlap) in Montana and Idaho, and compare these patterns to the general patterns predicted by our models from Step 1. We will then parameterize the models with data for Montana and Idaho and generate specific predictions of territorial behavior under each model. We will compare these predictions to actual locations of GPS-collared wolves in Montana and Idaho to test for concordance, and use multimodel inference to identify the models that most closely predict real territorial behavior. We will conduct sensitivity analyses and provide easy-to-use deliverables.

3. **Develop a suite of mechanistic group size models.** These models will capture the potential causal mechanisms we hypothesize structure social behavior of wolves. We will run simulations to provide general predictions of social behavior under each model.

4. **Identify the most predictive group size model for wolves in Montana and Idaho.** As with the territory models, we will test for concordance between model predictions and general patterns observed in real wolf packs. We will then parameterize the group size models with data for Montana and Idaho and generate specific predictions of social behavior under each model. We will compare these predictions to actual group sizes of wolves in Montana and Idaho as identified through monitoring data. We will use multimodel inference to identify the models that most closely predict actual group sizes. As with the territory models, we will conduct sensitivity analyses and provide easy-to-use deliverables.

5. **Develop empirical models for territory and group size.** We will compare the results from steps 1 – 4 to the empirical models we develop to identify the advantages and limitations of each approach.

6. **Calibrate estimates of abundance.** We will use our models for territory and group size alongside POM to calibrate estimates of abundance of wolves in Montana and Idaho. The models will enable region-specific predictions in territories and group sizes to improve abundance estimation. These deliverables will furthermore enable managers to predict the effects of management actions by adjusting inputs, e.g., to represent increased harvest pressure to predict how territories and pack sizes will change under different harvest levels.

**2.3 Progress**

Step 1 is complete, fulfilling our first deliverable for Project 1 on target (end of 2017). We present in this report a summary of 2 IBMs from the full set created. A full manuscript is in preparation.
2.4 Methods

Developing IBMs for Step 1 comprised 3 primary components:

1. **Establish a model framework.** Before building the models, we determined the general framework for their structure based on behavioral theory.
2. **Develop a suite of mechanistic territory models.** Each model included hypothesized causal mechanisms of territorial behavior.
3. **Run simulations and summarize results.** This allowed us to make general predictions useful for comparing to patterns in empirical observations.

2.4.1 Model framework

Our objective was to model how packs select annual territories to predict such characteristics as territory size, location, and overlap to calibrate POM. Accordingly, we aimed to model territory selection to represent the sum of a pack’s movements rather than the movements themselves. To model territory selection, the landscape can be represented as a continuous grid of patches which packs select to add to their territories (e.g., Fig 2.2). For each pack, the sum of patches selected is the territory, and the summary statistics of interest such measures as territory size and overlap.

We selected a mechanistic modeling framework to provide models predictive at any spatiotemporal scale and reduce future needs for monitoring wolves and collecting data. We designed the mechanistic models based on theory of how carnivores select territories. Carnivores are likely adapted to choose economic territories that maximize value, i.e., by maximizing benefits and minimizing costs of territory ownership (Darwin 1859, Brown 1964, Brown and Orians 1970, Emlen and Oring 1977, Krebs and Kacelnik 1991, Adams 2001). Like other carnivores, we also expect that wolves are adapted to defend the smallest territory possible that meets a threshold of resources for survival and reproduction (Mitchell and Powell 2004, 2007, 2012).

Building mechanistic territory models necessitated developing a set of hypotheses about which benefits and costs of territorial behavior are likely most fundamental to wolves. Conceivably, numerous benefits and costs could affect how patches are valued.

![Fig. 2.2. Example of a simulated landscape where packs have formed territories. Where patches have not yet been selected, bright green patches are of high prey benefit; yellow medium, and red low. Gray lines represent major roads. Black patches indicate overlapping territories.](image-url)
during territory selection. After extensive literature searches and consideration, we hypothesized that the causal mechanisms of territorial behavior include the benefits of prey and costs of travel, competition, and humans (Brown & Orians 1970; Adams 2001; Mitchell & Powell 2004, 2007, 2012). Food resources are required for all animals, and black bears (Ursus americanus) were shown to structure home ranges optimally with respect to the spatial distribution of food resources (Mitchell & Powell 2004, 2007, 2012). Lack of travel costs would imply that territories should be limitless in size because packs would travel any distance to reach a patch. Lack of competition costs would allow territories to overlap completely. Lastly, humans are an important source of mortality for wolves; their presence likely represents a key cost to territorial behavior.

2.4.2 Mechanistic territory models

Each competing model defined a specific hypothesis for how packs value patches for territories. Our set of models included combinations of hypotheses that wolves select territories based on the benefits of prey and costs of travel, competition, and humans. The 2 models we present here (out

Fig. 2.3. Example simulated landscapes where prey distribution ranges from evenly to highly clumped and prey abundance ranges from low to high. Human use also ranges from low to high. Landscapes are 200 × 200 patches in size, and no 2 landscapes are exactly alike. Patches were technically scale-less at this stage. In step 2 they will be set to represent actual spatial extents (e.g., 1, 5, or 10 km²) based on the resolution of available data.
of our larger subset) hypothesized that wolves select patches based on benefits of prey and A) costs of travel and competition, or B) costs of travel, competition, and humans. Model B differed by including cost of humans; we hypothesized this cost may have changed post-delisting with implementation of harvest. These models will allow us to investigate this possibility in Step 2.

We designed and tested the models in the program NetLogo (Wilensky 1999). The landscape was represented as a grid of 200 × 200 patches on which packs formed territories. Each patch was associated with a benefit of prey and costs of travel, competition, and humans. Our goal at this stage was to predict how territories would vary under different scenarios. Accordingly, within any given simulation, the landscape contained a particular prey distribution (evenly to highly clumped), prey abundance (low to high), and level of human use (low to high; Fig. 2.3). The simulations also enabled exploring how wolves would structure territories if they perceived competition and humans to have various levels of costs. Accordingly, we set the costs of competition and humans between low and high for any given simulation.

Following behavioral theory, packs acquired patches for annual territories as economically as possible (Fig. 2.4). One pack colonized the landscape at a time. The pack selected patches for its territory in order of value. Patch values were the benefit of prey in a patch discounted by the costs associated with the patch (competition and travel for Model A, and competition, travel, and humans in Model B). A simulation continued until all packs formed territories and there were insufficient resources to enable more packs to colonize.

2.4.3 Simulations

To learn about our models, we completed simulations and collected data on the results, e.g., each pack’s territory size and overlap. We ran 25 simulations for each combination of prey distribution, prey abundance, human use, cost of competition, and cost of humans. This yielded 675 simulations for Model A and 8100 for Model B (this higher value reflected the many combinations of human use and cost of humans).

We used program R (R Core Team 2014) to summarize results. Summaries included territory size (number of patches), territory overlap.
(percentage of territory patches shared with another pack), human avoidance (mean cost of humans in each patch in the territory minus the mean cost of humans in the landscape), and numbers of territories. We calculated mean results over each prey distribution, prey abundance, human use, colonization order, cost of competition, cost of humans, and model.

2.5 Results

Our simulations predicted patterns related to territory size, territory overlap, avoidance of humans, and number of territories, as follows.

2.5.1 Territory size

Territory size varied by prey distribution, prey abundance, and model (Fig. 2.5). Territories were larger in areas of low prey abundance and where prey were evenly clumped. If wolves ignored humans (Model A), territory size varied less and generally was smaller at comparable prey distributions and abundances than if wolves viewed humans as a cost (Model B). For highly clumped prey, however, territories were larger when wolves ignored human costs.

Territory size also varied somewhat with human use under Model B (Fig. 2.6). As human use increased from low to high, mean territory size increased when prey were evenly or moderately clumped, and decreased when prey were highly clumped.

![Mean territory size by model](image)

Fig. 2.5. Mean territory size decreased as prey became more clumped and as prey abundance increased. Territory sizes were larger for Model B (which includes cost of humans) except where prey were highly clumped.
Mean territory size varied by colonization order (Fig. 2.7). Later colonizers established larger territories. Where prey were highly clumped, earlier colonizers had among the smallest territories observed and later colonizers the largest. This pattern was strongest for low prey abundance.

2.5.2 Territory overlap

Mean overlap among territories was greater where prey were highly clumped and at high abundance (Fig. 2.8). Model A predicted greater overlap than Model B at comparable prey distributions and abundances, and predicted more overlap where prey were highly clumped. Mean overlap among territories depended on cost of competition (Fig. 2.9). Overlap quickly dropped to 0% as cost of competition increased.

2.5.3 Additional responses to humans

In addition to responses to humans noted above, responses to humans were measured as degree of human avoidance. Mean human avoidance varied by cost of humans and level of human use (Fig. 2.10). Because packs ignored cost of humans in Model A, they exhibited no avoidance. Under Model B, avoidance was greater when cost of humans was higher. As cost of humans increased from low to high, avoidance increased most drastically for high levels of human use.

2.5.4 Number of Territories

Numbers of territories varied by prey distribution, prey abundance, and model (Fig. 2.7). Territories were least numerous where prey abundance was low. More packs formed territories where prey were highly clumped. Fewer formed when packs considered human costs (Model B).

![Diagram](image_url)

Fig. 2.6. Mean territory size increased or decreased with higher levels of human use, depending on prey distribution.
Fig. 2.7. Mean territory size varied by colonization order (e.g., $1 = 1^{st}$ pack to select a territory). Late colonizers established larger territories, particularly where prey were highly clumped and at low abundance. Results also provided mean # of packs. Fewer packs formed territories on landscapes of lower prey abundances, and when they factored in human costs (Model B).

Fig. 2.8. Mean territory overlap was greatest where prey were highly clumped. Model A had a wider range of overlap across prey distributions and consistently greater overlap than Model B (which includes costs of humans).
Fig. 2.9. Mean territory overlap decreased as cost of competition increased, and varied by model and prey distribution.

Fig. 2.10. Mean human avoidance increased with increasing cost of humans and human use (negative values indicate greater avoidance). Human use of “none” indicated 0 costs associated with human use (e.g., Model A).
2.6 Discussion

A primary deliverable for this project is a suite of territory models that will be useful for calibrating POM. We have completed our suite of territory models and Step 1 of this project on target with the project timeline. The models predict and account for how territory size and overlap may vary across Montana and Idaho. Such predictions will be critical for calibrating POM estimates in later steps of this project. At this stage, our models allow us to make general predictions of patterns we may observe empirically; these predictions are particularly useful for Step 2. Below, we discuss how our models will help Montana and Idaho meet management needs. We outline our models’ general predictions for territory size, territory overlap, responses to humans, and numbers of territories; we also discuss example applications of our models’ general predictions. More details, models, and predictions will be presented in our manuscript about these models (in progress).

2.6.1 Territory size

Ability to predict territory size and its spatiotemporal variation is fundamental to calibrating POM estimates. Accordingly, our models allow us to predict territory size and account for how it may vary spatiotemporally across Montana and Idaho based on factors such as prey distribution, prey abundance, human use, and population size. POM currently relies on the assumption that average territory size is 600 km² statewide. Over- or under-estimating territory sizes will directly influence the number of packs predicted by POM. If territories are larger than 600 km², number of packs and overall abundance will be overestimated. If smaller than 600 km², estimates will be biased low. By accounting for variation in territory sizes rather than assuming a consistent territory size statewide, future POM estimates for number of packs and abundance of wolves will be more accurate and region-specific. Below, we discuss how prey distribution, prey abundance, human use, and population size affect predicted territory sizes.

Because distribution of prey may affect wolf territories (Fig. 2.5), our models will ensure territory sizes incorporated into POM remain calibrated across the spatially and temporally variable prey populations in Montana. The models demonstrate how prey distribution may affect territory size; assuming territory size is consistent regardless of prey distributions may thus over- and under-estimate abundance from POM in any given area. Our models predict territories to be, on average, larger in areas of Montana and Idaho where prey are more evenly clumped compared to more highly clumped. Importantly, these predictions are seasonal. Where ungulates are migratory, prey benefit of patches will shift seasonally, causing packs to adjust territories to the changing values of patches on the landscape. Once we parameterize the models with empirical data in Step 2, the simulations will account for seasonal changes in spatial distributions of ungulates. The sum of season-specific predictions will provide year-round territory predictions, which will likely be larger in areas where ungulates tend to be more migratory. As an example application of these predictions, we might expect seasonal territories to be larger in areas primarily occupied by deer (*Odocoileus* spp., e.g., northwest Montana) versus elk (*Cervus* spp., e.g., south-central Montana).
*canadensis*, e.g., southwest Montana), because deer tend to be more evenly spaced than large gregarious elk herds. Across the year, however, packs in southwest Montana may have larger territories if they respond to long distance elk and deer migrations. E.g., elk herds in the Yellowstone region may migrate 40 – 60 km (Nelson et al. 2012, Middleton et al. 2013), and mule deer (*O. hemionus*) may migrate 20 – 158 km (Sawyer et al. 2005). In contrast, in the rugged terrain of northwest Montana, white-tailed deer (*O. virginianus*) comprise the bulk of the ungulate population and generally exhibit shorter-distance elevational migrations. We would thus expect a more consistent prey distribution across seasons in northwest Montana. We expect that, after accounting for shifting prey availability, annual territories of wolves in northwest Montana will be smaller than those in southwest Montana.

Given that abundance of prey may also affect wolf territories (Fig. 2.5), our models will ensure territory sizes incorporated into POM remain calibrated across variable abundance of prey, which will further increase accuracy of POM estimates. The models demonstrate how territory size may vary based on prey abundance, e.g., territory sizes may be much larger in areas of low prey abundance compared to areas of high prey abundance. Accordingly, POM’s current assumption of a consistent territory size statewide may be overestimating number of packs in areas of low prey abundance, or underestimating number of packs in areas of high prey abundance. As an example application of this prediction, we might expect territory sizes to be larger in MFWP Region 5 than Region 3 where the ungulate populations differ by two-fold (~78,000 deer and elk in Region 5 versus ~146,000 in Region 3; fwp.mt.gov, accessed 2 Feb 2018). This may lead POM estimates in Region 5 to be biased high, and, conversely, estimates in Region 3 to be biased low.

Additionally, because human use may affect wolf territories (Fig. 2.6), our models will ensure territory sizes incorporated into POM remain calibrated across the spatially and temporally variable levels of human use in Montana. Our models demonstrate how territory size may vary across Montana and Idaho based on human use of the landscape. Specifically, when prey distribution is evenly or moderately clumped, Model B predicts slightly larger territories in areas with higher human use compared to areas of lower human use; conversely, where prey are highly clumped the model predicts the opposite (i.e., smaller territory sizes where human use is higher). As an example application of these predictions, when comparing territories in areas of Montana with high human use (e.g., close to cities) to areas of low human use (e.g., designated wilderness), we may expect to observe, on average, slightly larger territories where prey are evenly or moderately clumped, and slightly smaller territories where prey are highly clumped.

In Step 2, we will compare general predictions from each model to empirical patterns to ascertain model usefulness across spatiotemporal scales; Model B’s predictions for territory sizes will be particularly informative. We hypothesized that wolves will associate humans with higher costs post-delisting and with implementation of harvest. If our hypothesis is supported and Model B suitably captures this behavior, post-delisting we may observe: a) a greater range in territory
sizes; b) an increased mean territory size where prey distributions are evenly or moderately clumped; and c) a decreased mean territory size where prey are highly clumped (Fig. 2.5). We also hypothesized that wolves will associate humans with higher costs outside of protected areas. Accordingly, we might also expect to observe these patterns outside of Yellowstone National Park (YNP) compared to within the park.

Because wolf population size may also affect wolf territories (Fig. 2.7), our models will ensure territory sizes incorporated into POM remain calibrated across the spatially and temporally variable wolf populations in Montana. The models predict that the first packs to claim territories in an area may have smaller territories than their counterparts that colonize later. Average territory size may gradually increase as more packs form territories. Variation in territory sizes may similarly increase. Our models predict this pattern may be most noticeable in areas with highly clumped prey; where there are already many other packs, the newest packs may have among the largest territories observed. As an example application of this prediction, territories occupied for the longest in northwest Montana (e.g., some of those in the North Fork) may be among the smallest observed in that region. The same may be true for early packs in YNP. Furthermore, given the clumped nature of prey resources in YNP, the newest packs may, on average, have among the largest territories observed in Montana (if new packs do not simply usurp and maintain an old pack’s territory).

2.6.2 Territory overlap

Ability to predict territory overlap and its spatiotemporal variation is similarly critical for calibrating estimates from POM. As with territory size, our models allow us to predict and account for how territory overlap may vary spatiotemporally across Montana and Idaho. POM currently assumes overlap among territories is minimal and at consistent levels statewide. Over- or under-predicting overlap among territories will directly influence accuracy in the estimated numbers of packs from POM. I.e., where overlap among territories is greater than currently assumed, abundance may be underestimated, and where overlap is less than currently assumed, abundance may be overestimated.

Because territory overlap may be affected by the distribution and abundance of prey and level of human use (Fig. 2.8), our models will ensure territory overlap incorporated into POM remain calibrated across the spatially and temporally variable prey populations and levels of human use in Montana. Our models predict that territory overlap may be highest in areas where prey are more highly clumped and of higher abundance. Territory overlap is also predicted to be lower under Model B, demonstrating that if wolves perceive humans to be a cost to territory ownership, overlap may be lower. As an example application of these predictions, we might expect overlap to be greater in southwest Montana due to a highly clumped elk population (i.e., compared to deer, see above) and high abundance of ungulates. Additionally, if there is support for our hypothesis that wolves perceive humans as more costly post-delisting, we also may expect to see less overlap among territories today than pre-harvest.
We will further refine the predictive capacity of our models in Step 2 by investigating how wolves perceive the cost of competition; this will further calibrate predictions of territory overlap to increase accuracy of POM estimates. Our models demonstrate how overlap among territories depends on how wolves perceive cost of competition (Fig. 2.9). After parameterizing our models with empirical data in Step 2, we will determine which level of costs yields predictions that most closely match wolf territories. Real packs will therefore reveal the relative costs of competition compared to other benefits and costs of territorial behavior.

2.6.3 Additional responses to humans

Ability to predict how wolves will vary territorial behavior in response to human influences is useful in several ways for calibrating estimates from POM. As discussed above, our models allow us to account for how territory size and overlap may vary spatiotemporally across Montana and Idaho in response to humans, and this will directly calibrate POM. Two additional uses merit further discussion. First, the models predict how wolves may select territories to avoid humans. These predictions will be useful towards identifying the most appropriate models for calibrating POM. Secondly, ability to predict responses to human influences means our models will be useful for predicting the effects of management actions. We address these two uses below.

In Step 2, we will identify the most appropriate model for each area of Montana and Idaho to calibrate POM predictions. The degree to which territories avoid humans will be particularly useful for identifying which models better capture territorial behavior in each area (Fig. 2.10). The models demonstrate that if wolves perceive humans as a cost to territory ownership (Model B), territories will show avoidance of humans, otherwise they will show no response (Model A). Additionally, where human use is higher, Model B predicts that territories will be selected in areas that better minimize exposure to people. Where these predictions match empirical observations, Model B will be the more appropriate model for calibrating POM; elsewhere, Model A may be the more appropriate model. For example, Model A may suitably predict territories of wolves within YNP where cost of humans may be less important, whereas Model B may better predict territories in more urban areas of Montana.

Also in Step 2, we will refine the predictive capacity of our models by investigating how wolves perceive the cost of humans; this will further calibrate model predictions to increase accuracy of POM estimates. Our models predict that avoidance of humans will be stronger if wolves associate humans with higher costs (Fig. 2.10). Once we parameterize models with empirical data and compare predictions of human avoidance to empirical observations, real packs will reveal the relative costs of humans compared to other benefits and costs of territorial behavior.

Our models will also be useful for predicting the effects of management actions. This will directly assist management decision-making and integrate well with the adaptive harvest management component of Project 2. E.g., managers will be able to adjust model components to understand how various levels of human pressure (e.g., to represent altered hunting pressure) will
affect human avoidance, territory sizes, etc. Managers will also be able to predict how the removal of any given pack (i.e., through depredation removals) may affect other packs.

2.6.4 Number of territories

Though not an original deliverable, ability to predict numbers of territories and how this varies spatiotemporally can also be useful within the POM framework. Our models predict how number of territories may vary by prey distribution, prey abundance, and human use. We could use predictions for numbers of territories in two ways. First, we could incorporate these predictions within POM to calibrate estimates of colonization and resulting abundance; e.g., new colonization may be less likely in areas near predicted capacity. Secondly, we could compare our models’ predictions to number of packs estimated by POM as an indicator of accuracy in POM predictions.

Given these potential uses, our models’ ability to predict number of territories may be useful for POM. Our models predict slightly fewer packs in areas with evenly dispersed prey, and far fewer packs in areas with low prey abundance (Fig. 2.7). Additionally, our models predict somewhat fewer packs under Model B, if wolves perceive humans to be a cost of territory ownership. As an example application of these predictions, we may expect fewer packs in MFWP Region 5 than in Region 3 where prey abundance differs two-fold. We might also expect to see fewer packs post-delisting or outside of protected areas, if wolves associate humans with higher costs than they did pre-harvest.

2.6.5 Ongoing work

Our next step will fulfill the final territory model deliverable (due late 2019) by identifying the most predictive models from the full suite of models we have developed. We are currently preparing to formally summarize general patterns of observed territories in Montana and Idaho for Step 2. General concordance between empirical observations and model predictions (e.g., including those discussed above) will indicate that the models adequately capture mechanisms of territorial behavior. We are also preparing to parameterize the models with empirical data. Running simulations with empirical data will allow us to generate specific predictions of territorial behavior for wolves in Montana and Idaho. We will compare these predictions to territories of GPS-collared wolves to investigate which models most closely predict territorial behavior of wolves in Montana and Idaho.

Our final territory models will provide spatially-explicit predictions of territory size and overlap to calibrate POM, as discussed above. Furthermore, upon completing the territory models and parameterizing them with empirical data, they will also be useful for predicting locations of territories. We can use this feature, for example, to further design a finely detailed, spatially-explicit grid for POM to replace the current 600 km² grid and further increase accuracy of abundance estimates. We can also use this feature to predict locations of future territories (e.g.,
in currently-unpopulated areas of central Montana), or the effects of removals of packs (e.g., through depredation removals).

After identifying the best models, we will determine sensitivity to model inputs and level of data required for future use. This will demonstrate model robustness and the minimum data that will be required in Montana and Idaho to calculate accurate estimates of abundance in POM. Some model components will largely arise from the model itself (i.e., competition) or be easily measured using existing, widely-available data (i.e., travel costs that are based on Euclidean distance). Other inputs (i.e., prey and humans) will use basic sub-models that we will design to require existing data and little updating. More details will follow in subsequent manuscripts and reports.

ACKNOWLEDGEMENTS

This project would not be possible without the generous assistance of biologists and managers at Montana Fish, Wildlife and Parks, including Abby Nelson, Ty Smucker, Kent Laudon, Tyler Parks, Nathan Lance, Mike Ross, Diane Boyd, Molly Parks, Brady Dunne, Liz Bradley, Jessy Coltrane, Kelly Proffitt, John Vore, Quentin Kujala, Neil Anderson, and Bob Inman. Biologists and managers at Idaho Fish and Game also provide generous assistance, including David Ausband and Mark Hurley. We also thank landowners for allowing access for trapping and collaring efforts. Additionally, faculty and staff at the University of Montana provide invaluable support.

LITERATURE CITED


